

基于多特征联合学习的脓毒症死亡风险预测模型

文婷¹, 余雷², 李腊全¹

(1. 重庆邮电大学数学与统计学院, 重庆 400065; 2. 重庆医科大学第二附属医院急诊科, 重庆 400010)

摘要: 脓毒症是一种由感染而引起的全身炎症反应综合症, 在重症监护病房中具有较高的死亡率。然而, 现有的预测方法通常依赖于从特征集合中提取单一特征子集, 未能充分利用特征子集之间的复杂关联性, 即结构互信息, 从而限制了预测准确性。结构互信息不仅衡量了同一粒度下特征之间的依赖性, 还揭示了不同粒度下特征之间的复杂关系, 使其能够更精确地捕捉病情的细微变化。为了解决这一问题, 本文提出了一种新的脓毒症预后模型, 通过深入挖掘电子病历中的结构互信息, 以显著提高死亡风险预测的准确性。实验结果表明, 本文的预后模型在预测准确性方面表现出显著优势, 为临床医生提供了更可靠的死亡风险评估和明确的决策支持。

关键词: 深度学习; 集成学习; 特征学习; 脓毒症预后; 决策支持

中图分类号: TP391.4 **文献标志码:** A

引用格式: 文婷, 余雷, 李腊全. 基于多特征联合学习的脓毒症死亡风险预测模型[J]. 数据采集与处理, 2026, 41(3): 869-881. WEN Ting, YU Lei, LI Laquan. A sepsis mortality risk prediction model based on multi-feature federated learning[J]. Journal of Data Acquisition and Processing, 2026, 41(3): 869-881.

引言

脓毒症是一种严峻的全球健康挑战, 总体死亡率超过30%, 在重症监护病房(Intensive care unit, ICU)中的死亡率尤为突出^[1-2]。最新发布的“第三届脓毒症和脓毒性休克国际共识定义”强调了脓毒症感染与多器官衰竭之间的密切关系, 并明确指出脓毒症的管理和预后仍是全球卫生研究的重要方向^[3-4]。尽管近年来在脓毒症诊断和治疗技术方面有所突破, 但研究表明, 早期识别高风险脓毒症患者仍然是进一步改善预后的关键^[5]。然而, 由于脓毒症定义的不明确性、感染来源的难以预测性以及病情进展的快速性, 早期准确预测脓毒症患者的死亡风险依然是一项重大挑战, 这给临床医生和研究人员带来了诸多困难。首先, 目前广泛使用的临床评分方法受限于传统医学专家的经验。其次, 预先设定的常规指标通常较为有限和单一, 难以应对复杂的临床场景。在现代临床诊断中, 如何有效地识别和利用关键特征成为亟待解决的问题^[6]。此外, 现有基于生理参数的临床评分方法在应对复杂情况时缺乏灵活性, 限制了其预测能力的进一步提升。最后, 由于脓毒症发病迅速, 患者的个体临床数据积累有限, 使得准确预测脓毒症患者的死亡风险更加困难^[7]。

为了应对这些挑战, 研究人员积极探索自适应特征选择算法, 以提升脓毒症患者死亡风险评估的准确性。Dias等^[8]使用多变量分析发现, 从普通病房转入ICU的无发热脓毒症患者的死亡率显著高于发热患者, 尽管这一方法在特定情境下有效, 但仍受限于模型的复杂性和原始数据的分布特性, 难以获

得具有广泛适应性的通用特征。神经网络由于其强大的拟合能力,可以帮助模型聚焦于脓毒症相关指标中的关键特征,在提升脓症患者死亡风险预测的准确性方面表现出巨大潜力。Shah等^[9]开发了一种用于早期预测脓毒症的深度学习模型,验证了神经网络在实际临床环境中的可行性。Lauritsen^[10]等提出了一种用于早期检测脓毒症的深度学习系统,以自动从原始事件序列中学习关键因素及其相互作用的表示,减少了人工特征设计的需求。这些研究成果凸显了数据挖掘在脓毒症风险评估中的重要性,尤其是在提取关键特征方面,显著提升了预测的准确性。然而,现有的深度学习方法大多依赖简单的自适应特征选择,未能充分挖掘特征及其子集之间复杂的相互依赖性,即结构互信息。由于脓症患者病情变化复杂,各种生理特征之间可能存在非线性和多层次的相互作用,单一模型难以有效地捕捉这种复杂的关联性。这一局限性削弱了现有模型在捕捉细微病情变化和提升预测精度方面的能力,使得准确评估脓毒症患者的死亡风险更加困难。因此,如何有效捕捉特征之间的复杂依赖关系,构建能够融合关键特征及其交互信息的预测模型,成为当前亟需解决的重要问题。

针对这一问题,本文提出了一种新的脓毒症预后预测模型(Sepsis prognosis prediction model, SPPM),旨在通过深入挖掘电子病历(Electronic medical record, EMR)数据中的结构互信息,显著提升脓症患者死亡风险预测的准确性和实用性。首先,在国际公开的脓症患者EMR数据集上,采用多特征选择策略进行数据划分,并基于联邦学习(Federated learning, FL)算法集成多层感知机(Multilayer perceptron, MLP)的结果。多特征选择策略遵循临床预测中广泛应用的算法^[11],划分出多个特征子集。针对每个特征子集,本文独立训练了一个MLP模型。由于这些数据子集在患者身份标识符上完全重叠,而在特征维度上部分重叠,这种策略使每个MLP模型能专注于特定特征子集,从而增强模型对特定临床特征的敏感性。最终,通过FL算法对这些独立训练的MLP模型进行集成,实现了在不共享原始患者数据的情况下的协同训练,显著提升了模型的泛化能力和预测性能。

在重症监护医学信息库(Medical information mart for intensive care, MIMIC-IV 2.2)^[12]这个国际公开数据集中,SPPM模型多次实验结果显示,在评估脓症患者ICU入住后30天内的死亡风险方面,SPPM模型的性能优于现有主流机器学习模型,达到了临床辅助诊断的标准。此研究不仅为脓毒症的临床预测提供了新视角,还为医疗数据隐私保护和模型集成提供了创新性的解决方案。图1所示的模型架构详细描述了本文方法。

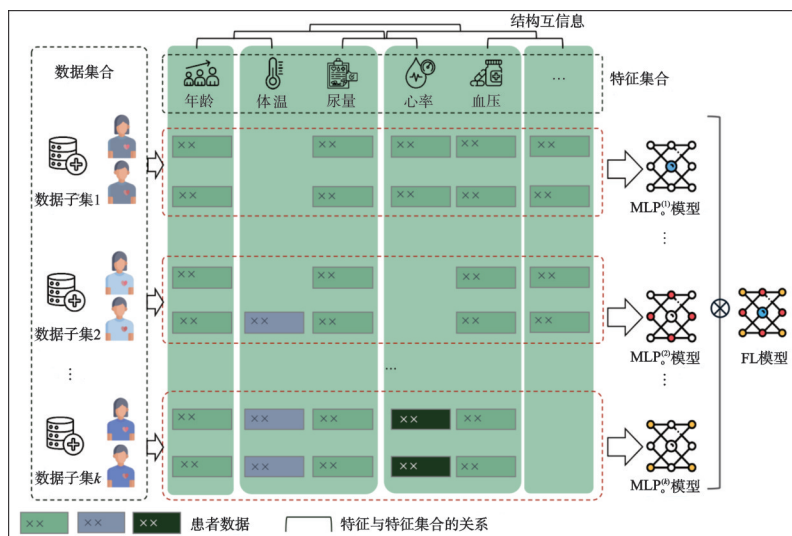


图1 基于多层感知机和联邦学习的脓毒症预后模型架构

Fig.1 Sepsis prognosis model architecture based on multilayer perceptron and federated learning

1 相关工作

在医学领域,脓毒症的临床预测模型在提高ICU患者短期死亡风险预测准确性方面发挥着至关重要的作用。随着人工智能技术的发展,基于机器学习和深度学习的方法在预测精度和个性化风险评估方面展现出一定优势。相较于传统统计学方法,这类模型能够更充分地利用多维临床数据,从而提升脓毒症风险评估的准确性和适用性。这一趋势表明,人工智能驱动的临床预测模型正在成为改善患者预后、优化治疗策略的关键工具,在医学领域中发挥着越来越重要的作用。

特征选择是提升临床预测模型性能的关键步骤,其目的是在复杂的医疗数据中识别出模型对结局目标预测最具影响力的特征^[13]。通过特征选择,模型可以剔除冗余和无关信息,从而提高准确性和稳定性,同时提升训练效率。经典的特征选择方法包括包装算法、过滤算法和嵌入式算法。包装算法如Boruta算法^[14],通过将真实特征与伪特征(即人工生成的无信息特征)进行比较来优化特征集合,从而确定特征的重要性。过滤算法^[15]如信息增益(Information gain, IG)、随机森林(Random forest, RF)、对称不确定性(Symmetrical uncertainty, SU)以及Relief算法,依靠统计测度来评估特征的重要性。嵌入式算法则在模型训练过程中,通过参数优化来实现特征选择,典型算法包括最小绝对收缩与选择算子(Least absolute shrinkage and selection operator, LASSO)^[16]、递归特征消除(Recursive feature elimination, RFE)^[17]以及极致梯度提升(Extreme gradient boosting, XGBoost)^[18]。其中,XGBoost算法通过每次迭代调整特征重要性得分,自动选择关键特征,从而显著提升模型的预测性能和效率。

深度学习(Deep learning, DL)模型^[19-20],尤其是多层感知机(Multilayer perceptron, MLP)^[21-22],在处理复杂医疗数据和进行临床预测方面展现了强大的能力,已经在医学领域得到了广泛应用。MLP通过其多层结构,能够有效应对数据中的非线性关系,具备出色的非线性建模能力,这使得MLP在从电子健康记录中提取关键信息方面表现优异,帮助医生更准确地进行风险预测和制定个性化治疗策略。然而,MLP等深度学习模型的训练通常需要大量的数据,这可能会涉及到隐私保护和数据安全的问题^[23-24]。在提升模型性能的同时保障数据隐私,联邦学习(Federated learning, FL)^[25]与MLP的结合可能成为一种有效的策略。FL允许多个参与方在不用共享原始数据的情况下协同训练模型,从而避免了集中存储完整医疗数据带来的隐私风险^[26]。通过共享模型参数而非非数据,将不同数据子集训练的MLP模型整合在一起,有助于提升本文所提出的SPPM模型的预测性能。

综上所述,多层感知机与联邦学习的结合在脓毒症临床预测领域展现出了广阔的应用前景。现有研究通过引入MLP的非线性建模能力及FL在数据隐私保护方面的优势,为复杂医疗数据的处理和高风险患者的识别提供了有效策略。这些方法不仅提升了模型的预测精度,还降低了数据集中存储所带来的隐私风险。

2 脓毒症预后模型

2.1 数据来源

本文研究的数据来源于MIMIC-IV 2.2,为了访问这一数据库,本文一位作者获得了必要的认证(认证号:55376203),随后提取了研究相关的特征。由于该数据库中的患者健康信息是匿名的,因此不需要患者个人同意^[27]。根据Sepsis-3标准^[3],患者在入住ICU的最初24 h内被诊断为脓毒症满足如下条件:(1)确诊或疑似感染;(2)连续器官功能衰竭评估评分为两分或两分以上。本文排除了以下3种患者情况:(1)多次入住ICU和多次住院;(2)ICU住院时间少于24 h;(3)18岁以下的脓毒症患者。

2.2 基于联邦学习的多特征选择策略

特征选择和合理的数据划分在提升模型性能和解释性方面扮演了关键角色。针对脓毒症的临床预测,本文强调了基于特征相关性的数据划分方法的重要性。将高度相关的特征聚集于同一子集,可以减少数据子集间的协作需求,优化计算效率的同时进一步增强数据隐私保护效果,垂直联邦学习^[28]正是这种方法的重要应用形式。

在预测脓毒症患者入住ICU后30天内的死亡风险时,将与结局高度相关的特征划分到同一子集中,不仅显著提升了模型的训练效率,还降低了对跨子集数据整合的需求,从而有效提升了隐私保护效果。为了实现这种高效的数据划分,本文采用了多种特征选择方法,包括Boruta、IG、SU、RF、Relief、LASSO、RFE以及XGBoost算法等。这些方法通过不同的计算机制对特征的重要性进行综合评估,从而确保模型能够捕捉到对预测结果最具影响的特征。

本文选择上述特征选择算法,是基于在多样性、理论支持及实证效果方面的坚实基础。这些算法覆盖了统计度量、熵测量和树模型等不同类型的方法,以确保特征选择过程的全面性和鲁棒性。同时,这些算法在现有文献中被广泛应用,已通过多项研究验证其有效性。通过融合多种特征选择策略,本文旨在构建一个全面的特征评估框架,最大限度地提升模型的预测性能和稳定性,从而为临床决策提供更为丰富和可靠的支持。

设数据集 $D = \{(x_i, y_i)\}_{i=1}^I$,基于多个特征选择算法,生成多个特征子集,有

$$S_k = \{x_{ij} | j \in 1, 2, \dots, d\} \quad (1)$$

式中: I 表示患者总数, x_i 表示第*i*个患者的特征向量, y_i 表示患者的标签, S_k 表示特征向量的子集组成的第*k*个特征子集, x_{ij} 表示医学特征子集 S_k 中第*i*个患者的第*j*个医学特征, d 表示数据集 D 的维度。数据集 D 被划分为多个数据子集,具体来说,生成*k*个特征子集,其中*k*=8。通过这种方式,每个子集都有助于保护数据隐私,同时提高联邦学习的效率。

总的来说,基于多特征选择策略将数据集合理地划分为多个特征子集,不仅保护了数据隐私,还提高了模型的效率,这促进了跨模型训练,实现了更高效、更安全的临床预测模型构建。

2.3 基于关联注意力的深度学习模型

本文采用多特征选择策略,将脓毒症患者数据集划分为若干特征子集。在特征选择的基础上,进一步引入基于关联注意力的深度学习模型^[29],以提升预测效果。具体而言,为每个数据子集分别定制MLP模型,并对各自持有的特征进行独立训练。为了提高模型对重要特征的关注,本文在每个MLP模型中引入了关联注意力机制。该机制通过动态调整不同特征的权重,突出关键信息。具体来说,关联注意力机制根据每个特征与预测结果的相关性来分配权重,从而在训练过程中为重要特征赋予更高的权重。经此机制处理的特征子集用于训练独立的MLP模型,确保每个模型最大限度地利用其特征子集中的有效信息。在此之后,不再需要进行进一步的特征选择。关联注意力机制不仅简化了特征选择过程,还充分利用了每个特征子集中的信息,提高了模型的预测精度和计算效率。

每个特征子集 S_k ,分别训练一个定制化的MLP₀^(k),有

$$\text{MLP}_0^{(k)}: x_{ij} \rightarrow \hat{y}_k \quad (2)$$

式中 \hat{y}_k 为模型MLP₀^(k)在类别*m*的预测值。

于是,MLP₀^(k)模型的前向传播过程为

$$\begin{cases} \text{Sig}_k^{(1)} = \sigma(W_k^{(1)} X_k' + p_k^{(1)}) \\ \text{Sig}_k^{(2)} = \sigma(W_k^{(2)} \text{Sig}_k^{(1)} + p_k^{(2)}) \\ \vdots \\ \hat{y}_k = \sigma(W_k^{(L)} \text{Sig}_k^{(L-1)} + p_k^{(L)}) \end{cases} \quad (3)$$

式中: $\text{Sig}_k^{(L)}$ 为第 L 层的激活值, σ 为激活函数, $\mathbf{W}_k^{(L)}$ 为 L 层的权重矩阵, $\mathbf{p}_k^{(L)}$ 为相对应层数的偏置向量。

为了最小化损失函数, 对于第 k 个模型, 其参数更新公式为

$$\zeta_k \leftarrow \zeta_k - \eta \nabla_{\zeta_k} \Gamma(\zeta_k) \quad (4)$$

式中: η 表示学习率, $\eta \nabla_{\zeta_k} \Gamma(\zeta_k)$ 表示该损失函数对模型参数的梯度。

将新的特征作为输入, 基于 FL 集成 MLP 模型的梯度具体计算过程包括: 输出层的误差、逐层反向传播计算误差和每层的梯度, 分别用数学符号表示为

$$\begin{cases} \delta_k^{(L)} = \hat{y}_{k,i} - y_i \\ \delta_k^{(l)} = (\mathbf{W}_k^{(l+1)})^T \delta_{k,i}^{(l+1)} \odot \sigma'(a_{k,i}^{(l)}) \\ \nabla_{\mathbf{W}_k^{(l)}} \Gamma(\zeta_k) = \frac{1}{B} \sum_{i=1}^B \delta_{k,i}^{(l)} (\text{Sig}_{k,i}^{(l-1)})^T \\ \nabla_{\mathbf{p}_k^{(l)}} \Gamma(\zeta_k) = \frac{1}{B} \sum_{i=1}^B \delta_{k,i}^{(l)} \end{cases} \quad (5)$$

式中: $\delta_{k,i}^{(l)}$ 表示第 k 个 MLP 在输出层 L 对第 i 个样本的误差项, $\hat{y}_{k,i}$ 表示第 i 个样本的预测值, y_i 表示第 i 个样本的真实值, “ \odot ” 表示逐元素乘法, $\sigma'(\cdot)$ 表示激活函数对线性输入的导数, T 表示转置操作, B 为训练时批量大小, $a_{k,i}^{(l)} = \mathbf{W}_k^{(l)} \text{Sig}_{k,i}^{(l-1)} + \mathbf{p}_k^{(l)}$ 表示第 k 个 MLP 在第 l 层对第 i 个样本的线性输入, $\text{Sig}_{k,i}^{(l-1)}$ 表示第 k 个 MLP 在 $l-1$ 层对第 i 个样本的激活输出, $\nabla_{(\cdot)}$ 表示求梯度, $\Gamma(\zeta_k)$ 是第 k 个本地模型关于参数集合 ζ_k 的损失函数。

MLP 模型因其强大的拟合能力, 可以处理多种分类任务。引入关联注意力机制不仅简化了模型结构, 还进一步发挥了 MLP 的拟合能力。在 FL 框架下, 这一方法实现了在保护数据隐私的前提下构建高效、准确的临床预测模型, 为脓毒症的早期检测和干预提供了有力支持。通过垂直划分数据并进行协同计算, 各医疗机构之间可以更有效地协作, 共同提升医疗预测模型的性能。这种整合方法有助于在分布式环境中优化资源利用, 促进跨机构数据共享与合作, 同时保证患者数据的安全性和隐私性。

2.4 基于多层感知机和联邦学习的集成模型

基于多特征选择策略, 将原始数据集划分为多个特征子集, 并为每个特征子集定制化训练 1 个 MLP 模型。每个 MLP 模型独立处理其对应的特征子集并输出相应的预测值。具体来说, k 个特征子集 S_1, S_2, \dots, S_k , 分别训练 k 个 MLP 模型 $\text{MLP}_o^{(1)}, \text{MLP}_o^{(2)}, \dots, \text{MLP}_o^{(k)}$, 每个模型获得 1 个预测值 $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_k$ 。

在使用 MLP 模型进行训练时, 基于信息熵来衡量特征的重要性。此度量的目的是通过计算每个特征对结局的不确定性贡献, 识别出对模型性能影响最大的特征。具体而言, 通过信息熵来度量特征选择后所获得的信息增益, 从而确定各个特征的重要性排序, 揭示了各特征在预测结局中所占的相对比重。

给定概率分布 $p = (p_1, p_2, \dots, p_c)$, 其信息熵被定义为

$$E(p) = - \sum_{m=1}^M p_m \log p_m \quad (6)$$

式中: m 为类别的数量, p_m 为类别 m 的概率。

通过计算移除特征 x_{ij} 后的熵变化来衡量其贡献度。设 $\hat{y}_{i,-j}$ 为移除特征 x_{ij} 后的预测值, 则特征 x_{ij} 的贡献度 $\Delta E(x_{ij})$ 为

$$\Delta E(x_{ij}) = E(\hat{y}_i) - E(\hat{y}_{i,-j}) \quad (7)$$

上述经过信息熵评估的重要性特征被用于优化后续联邦学习模型的参数和架构。在联邦学习中, 不同节点 (如不同数据子集) 可能持有不同的数据分布和特征集, 分别训练独立的 MLP 模型。基于信息熵评估得到的特征重要性, 进一步优化联邦学习模型的参数与结构设计, 从而增强 SPPM 模型在异构数据分布下的鲁棒性与泛化能力, 确保了在不同数据分布下模型性能的稳定性。

在FL中,每个特征子集对应1个MLP模型,独立训练模型,再进行聚合。假设第 k 个数据子集的模型参数为 ς_k ,参数聚合使用加权平均,即

$$\varsigma_g = \sum_{k=1}^K \frac{N_k}{I} \varsigma_k \tag{8}$$

式中: N_k 为第 k 个数据子集的数据量, I 为所有数据子集的数据总量。

在模型集成阶段,本文采用了加权平均技术,对多个MLP模型的预测值进行聚合。具体来说,信息熵所确定的特征重要性权重,在联邦学习中的模型聚合过程中发挥了关键作用。这些权重指导了在不同节点上信息传递的优先级,确保了重要性较高的特征在加权过程中获得更大的比重。

特征子集被输入独立的模型中,每一个模型分别获得1个预测值。与此同时,去除弱模型结果,即剔除性能明显低于平均水平的模型,避免它们影响最终结果。由剩余 m 个预测值组成1个新的特征矩阵,有

$$Z = \begin{bmatrix} \hat{y}_1^{(1)} & \hat{y}_1^{(2)} & \cdots & \hat{y}_1^{(k)} \\ \hat{y}_2^{(1)} & \hat{y}_2^{(2)} & \cdots & \hat{y}_2^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{y}_i^{(1)} & \hat{y}_i^{(2)} & \cdots & \hat{y}_i^{(k)} \end{bmatrix} \tag{9}$$

式中 $\hat{y}_i^{(k)}$ 为第 m 个数据子集的第 i 个特征对整体 I 的预测值。

获得上述独立的预测值后,将其作为新的特征输入,形成1个新的特征矩阵。这个矩阵将所有模型的输出整合在一起,通过投票机制为进一步的决策提供信息。每个特征子集强调数据的不同维度,所训练的模型捕获数据的不同特性。基于FL框架,将 m 个MLP^(k)模型结果进行综合,于是有

$$\hat{y} = F_{\varsigma_g}(Z) = \text{mode}(z_1, z_2, \dots, z_m) \tag{10}$$

式中 F_{ς_g} 为 m 个预测值投票的最终预测函数。

集成多个模型,充分利用各个模型的优势,将多个模型的预测值作为输入,通过集成模型整合这些结果,从而获得最终的预测结果。这种方法在提升模型性能的同时,确保了数据的隐私和安全。一般方法与SPPM模型的对比如图2所示。具体算法步骤如算法1所示。

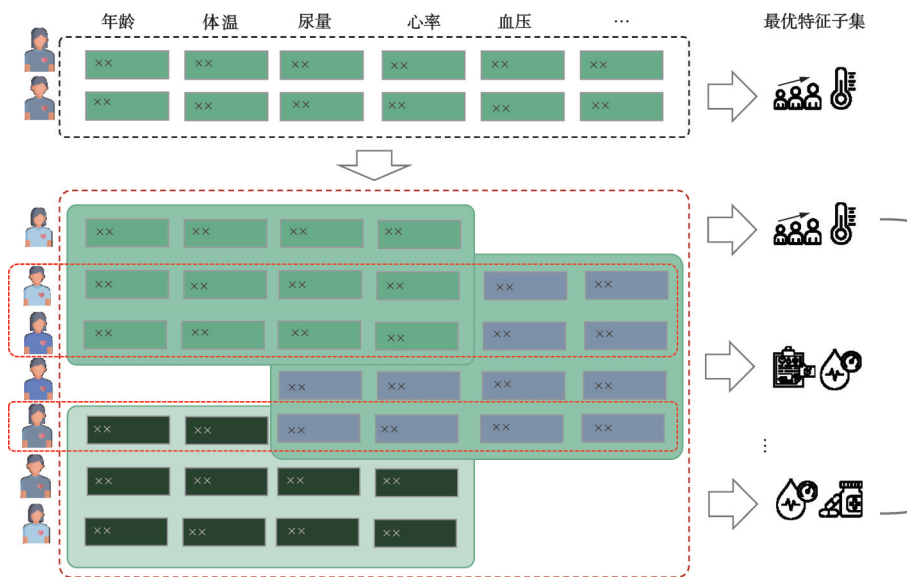


图2 从传统特征选择到SPPM模型的多特征学习策略

Fig.2 Multi-subset learning strategy from traditional feature selection to SPPM model

算法 1 SPPM 算法

输入:数据集 $D = \{(x_i, y_i)\}_{i=1}^I$, 样本总数 I , 特征向量 x_i , 标签 y_i ; 输出: 预测结果 \hat{y}

(1) 数据加载与预处理

(2) 特征选择和数据划分, S_1, S_2, \dots, S_k

(3) 训练 MLP 模型 $MLP_o^{(k)}: x_{ij} \rightarrow \hat{y}_k$

计算基于特征 x_{ij} 的贡献度 $\Delta E(x_{ij}) = E(\hat{y}_i) - E(\hat{y}_{i,-j})$, 根据贡献度排序

(4) 关联注意力机制: 基于贡献度大小分别对模型 $MLP_o^{(k)}$ 进行赋权, 使其更专注于重要特征, $\varsigma_g =$

$$\sum_{k=1}^K \frac{N_k}{I} \varsigma_k$$

返回 k 个 $MLP_o^{(k)}$ 模型的预测值 $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_k$

(5) 联邦学习集成: $Z = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_m]^T$

(6) 返回最终预测结果: $\hat{y} = F_{\varsigma_g}(Z) = \text{mode}(z_1, z_2, \dots, z_m)$

FL 的网络学习框架在保护患者隐私方面发挥了关键作用。该方法允许各参与方在本地保留患者数据, 仅共享或更新模型参数, 而非原始数据, 从而有效防止敏感信息的泄露, 保障了患者隐私的安全。此外, FL 通过分布式数据训练, 使模型能够在本地进行训练, 无需将所有数据集中处理。此方法不仅增强了模型的泛化能力和鲁棒性, 同时保持了数据的本地性, 避免了数据集中存储可能带来的安全风险。这些措施使得本文在保护患者隐私的同时, 实现了高效且安全的数据利用, 有助于推动医学研究的进步。

2.5 统计分析

本文根据患者因脓毒症入住 ICU 30 天内是否死亡, 将患者分为两组。此外, 正态和非正态分布的连续变量显示为四分数, 分类变量显示为实例数(%)。支持向量机(Support vector machine, SVM)被用于异常值检测, 并使用特征均值对检测到的异常值进行替代。对于特征缺失率低于 30% 的数据, 使用 R 软件包中的多重插补(Multivariate imputation by chained equations, MICE)方法进行了 50 次多重插补处理, 而缺失率大于或等于 30% 的特征则被剔除。在特征选择阶段, 本文使用 RFE 和 LASSO 回归等算法。假设特征数为 m , RFE 的计算复杂度为 $O(m^2)$, LASSO 的复杂度为 $O(n \cdot m^2)$ 。当 m 较大时, 这些特征选择方法的复杂度可能较高。然而, 通过降维显著减少了后续模型训练所需的计算量, 从而有效缓解了复杂度问题。此外, 本文还采用了 SMOTE-Tomek 采样方法来校正数据不平衡问题。所有的数据处理和分析均在 Python(Ver 3.11.4)和 R(Ver 4.3.1)环境中进行。

3 实验设置与结果**3.1 数据集**

本文从 MIMIC-IV 数据库的 73 181 名入住重症监护室的重症患者中筛选出 43 477 名脓症患者。其中, 通过排除非首次入住重症监护室的患者、入住重症监护室时间不足 24 h 且年龄在 18 岁以下的患者, 共筛选出 24 988 名脓症患者。最后, 共有 18 494 名具有 72 个相应特征的患者病例符合纳入研究的条件。其中, 3 502 例患者在 30 天内死亡, 14 992 例患者在 30 天内存活。此外, 本文使用 pgAdmin for PostgreSQL(V122.1)和 Navicat Premium(V12.0.28)提取了住院脓毒症患者的原始数据。脓毒症患者的临床变量包括 ICU 收治的脓毒症患者的人口统计学特征、疾病严重程度评分、实验室检查指标和其他临床特征。

3.2 实验设置

3.2.1 评价指标

本文通过5种常用且有效的指标对模型性能进行评估,具体包括准确度(Accuracy, ACC)、受试者工作特征曲线下面积(Area under the receiver operating characteristic curve, AUROC)、召回率(Recall)、精确度(Precision)以及 F_1 评分(F_1 -score)。这些指标各自从不同角度衡量了模型的优劣,能够全面反映模型在脓毒症患者死亡风险预测中的表现。

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

式中:TP为真正例,TN为真负例,FP为假正例,FN为假负例。

$$AUROC = \int_0^1 TPRd(FPR) \quad (14)$$

式中:TPR为真阳性率,FPR假阳性率。

$$F_1\text{-score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (15)$$

式中:Precision为准确度,Recall为召回率。

3.2.2 参数设置

本文的实验环境运行在Windows操作系统上,采用Python 3.8编程语言,并基于PyTorch 2.0深度学习框架进行开发。在超参数的配置与调整中,隐藏层使用了ReLU作为激活函数,而输出层则选用了Sigmoid激活函数。参数优化方面,采用Adam优化算法,学习率设置为0.001,以保证模型能够高效收敛。在所有参与联邦学习的节点中,各个MLP模型的架构设置保持一致,包括相同的隐层数和每层的隐单元数量。与此同时,考虑到输入数据子集的规模和特性相似,统一架构设置有助于提升整体模型的训练效率和预测性能。此外,为了提升模型的稳健性和泛化能力,训练过程中引入了Dropout策略,以减少过拟合的风险。

3.3 实验结果及分析

本文所提出的SPPM模型针对脓毒症患者入住ICU后30天内的死亡风险进行了深入评估。模型结果如表1所示。SPPM在多个关键性能指标上分别显著优于未特征选择和使用单个特征选择方法的主流机器学习模型,包括决策树(Decision tree, DT)^[30]、SVM^[31]、XGBoost^[32]、梯度提升(Gradient boosting, GB)^[33]、RF^[34]、卷积神经网络(Convolutional neural network, CNN)^[35]以及MLP^[36]。表1中,FS为特征选择,No-FS为未特征选择,SFS为单个特征选择方法,MFS为多特征选择策略,MLP_B为MLP_Boruta,MLP_IG为MLP_Information Gain,MLP_SU为MLP_Symmetrical Uncertainty,MLP_RF为MLP_Random Forest,MLP_Re为MLP_Relief,MLP_La为MLP_LASSO,MLP_RFE为MLP_Recursive feature elimination,MLP_XG为MLP_XGBoost。

不同模型评估指标可视化如图3所示,SPPM在ACC上达到了0.969 0,这一性能显著高于未进行特征选择的DT、SVM和XGBoost等主流机器学习模型。在AUROC上,SPPM以0.994 0的高性能领先,显示出极高的区分能力。与此同时,在Precision、Recall和 F_1 -score上,SPPM分别取得了0.976 9、0.960 6和0.968 7的优异成绩,均优于其他模型。此外,基于多特征选择策略的SPPM模型比基于多单个特征选择方法的模型性能高,这进一步证实了特征选择在提升模型性能中的重要性。

这些结果不仅反映了SPPM在脓毒症死亡风险评估任务中的优越性能,也验证了本文所提出方法

表1 不同模型评估脓毒症患者入住ICU 30天内死亡风险的对比结果

Table 1 Comparison of different models in evaluating 30-day ICU mortality risk of sepsis patients

FS	模型	准确率	ROC曲线下面积	精确率	召回率	F_1 -评分
No-FS	DT	0.807 0	0.807 0	0.793 9	0.829 2	0.811 2
	SVM	0.847 7	0.922 5	0.839 0	0.860 5	0.849 6
	XGBoost	0.873 1	0.921 7	0.933 7	0.803 1	0.863 5
	GB	0.881 2	0.931 7	0.974 7	0.782 8	0.868 2
	RF	0.908 8	0.963 4	0.968 5	0.837 2	0.901 7
	CNN	0.893 4	0.959 8	0.884 1	0.911 6	0.897 6
	MLP	0.926 4	0.976 0	0.913 3	0.942 2	0.927 5
SFS	MLP_B	0.886 7	0.951 2	0.886 8	0.886 5	0.886 7
	MLP_IG	0.910 0	0.969 9	0.886 4	0.940 6	0.912 7
	MLP_SU	0.913 7	0.966 7	0.902 5	0.927 5	0.914 9
	MLP_RF	0.914 7	0.963 4	0.898 1	0.935 6	0.916 4
	MLP_Re	0.933 7	0.977 3	0.919 6	0.950 6	0.934 8
	MLP_La	0.932 0	0.973 6	0.920 7	0.945 6	0.933 0
	MLP_RFE	0.916 9	0.962 2	0.902 9	0.934 2	0.918 3
MFS	MLP_XG	0.937 7	0.980 1	0.917 8	0.951 6	0.939 2
MFS	SPPM(ours)	0.969 0	0.994 0	0.976 9	0.960 6	0.968 7

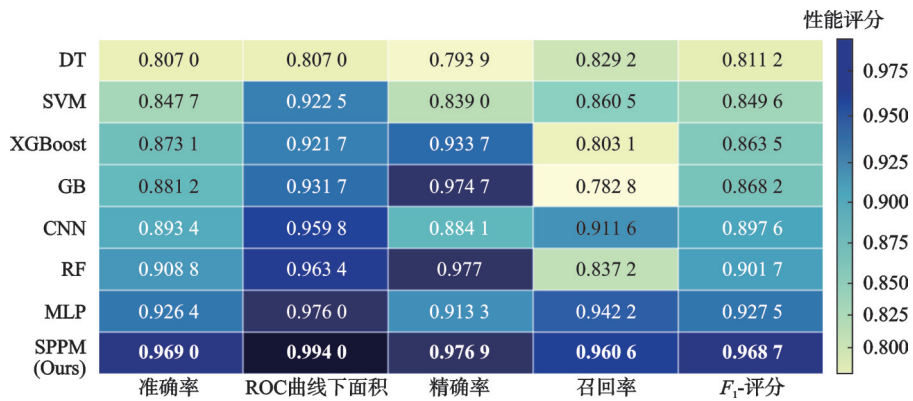


图3 不同模型的评估指标可视化

Fig.3 Visual comparison of evaluation metrics for different models

的有效性。SPPM基于FL方法,融合了多特征选择策略,在数据增强的同时保护了数据隐私。通过DL架构,该模型能够捕捉数据中的复杂非线性关系。

3.4 消融实验

本文通过消融实验评估了多特征选择策略对模型性能的影响。实验结果显示,特征选择在提高模型预测准确性方面具有重要作用。如表2和图4所示,当未进行特征选择时,MLP模型的ACC为0.926 4。然而,当采用单一特征选择方法(如XGBoost)后,MLP模型的准确率略有提升,达到0.937 7。这表明单一特征选择方法在捕捉脓毒症患者病情复杂性方面具有一定效力。然而,当采用多种特征选择方法融合的策略,即SPPM模型时,准确率显著提升至0.969 0。显著的性能提升表明,多特征选择方

表 2 基于多特征数据划分的消融实验结果

Table 2 Results of ablation experiments based on multi-feature data division

未特征选择	单特征选择	多特征选择	准确率	ROC 曲线下面积	精确率	召回率	F_1 -评分
✓			0.926 4	0.976 0	0.913 3	0.942 2	0.927 5
	✓		0.937 7	0.980 1	0.917 8	0.951 6	0.939 2
		✓	0.969 0	0.994 0	0.976 9	0.960 6	0.968 7

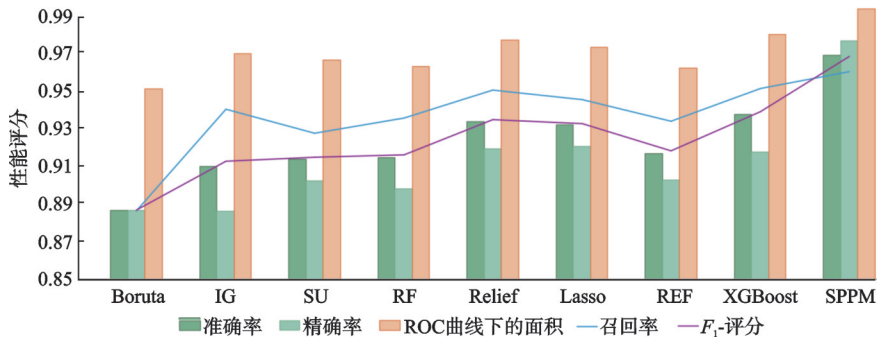


图 4 SPPM 与单一特征选择策略的性能比较

Fig.4 Performance comparison of SPPM with conventional feature selection methods

法在充分利用数据中的信息并增强模型预测能力方面具有显著潜力。SPPM 模型通过整合多种特征选择技术,提高了模型对特征重要性的识别能力。

在对比单一特征选择方法与 SPPM 模型的性能时,SPPM 模型在 ACC、AUROC、Precision、Recall 和 F_1 -score 等关键性能指标上均展现出其优越性。SPPM 模型的性能显著优于其他单一特征选择方法,这得益于其先进的多特征选择技术和集成模型训练策略。从表 2 中数据可以看出,未进行特征选择的模型在 AUROC 上达到了 0.976 0,而采用单特征选择后,AUROC 提升至 0.980 1,这进一步证实了特征选择在提升模型性能中的关键作用。SPPM 模型在 AUROC 上达到了 0.994 0,显示出极高的区分能力,这表明了其在预测脓毒症患者死亡风险方面的高准确性。

综上所述,消融实验结果充分证明了 SPPM 模型多特征选择策略在提升模型性能方面的显著潜力。该模型不仅通过优化特征组合有效提升了分类精度,还结合 MLP 和联邦学习进一步强化了整体性能。具体而言,优化特征组合显著提高了模型对患者死亡风险的敏感性,帮助医护人员更早发现高危患者并及时采取干预措施;分类精度的提升增强了预测的可靠性,降低了误诊和漏诊的可能性,从而减少资源浪费并避免延误治疗风险;整体性能的增强则显著提高了模型的临床适用性和普适性,使其能够应对复杂多样的患者数据。本文所提模型为脓毒症患者死亡风险的精准评估提供了有力工具,有望成为支持早期识别与临床决策的有效手段。

4 结束语

精准预测脓毒症患者 30 天内的死亡风险可有效帮助医生制定治疗策略,降低脓毒症在 ICU 中的高死亡率。现有方法多依赖单一的特征选择方法,未能充分利用特征子集间的复杂相关性,限制了预测准确性。本文提出的基于多层感知机和联邦学习的 SPPM 模型,通过将数据划分为多个特征子集,为每个子集训练定制化的 MLP 模型,再通过 FL 框架集成,提升了预测精度和泛化能力。实验结果表明,本文所提出的模型在预测准确性方面表现优异,为医生提供了更准确的死亡风险评估,辅助制定更有效的治疗计划,降低 ICU 脓毒症患者的死亡率,并减少医疗资源的消耗。未来可进一步优化模型参数和集成策略,以应对更多临床应用场景。

参 考 文 献:

- [1] SONG J, PARK D W, MOON S, et al. Diagnostic and prognostic value of interleukin-6, pentraxin 3, and procalcitonin levels among sepsis and septic shock patients: A prospective controlled study according to the Sepsis-3 definitions[J]. *BMC Infectious Diseases*, 2019, 19(1): 1-11.
- [2] LIU V, ESCOBAR G J, GREENE J D, et al. Hospital deaths in patients with sepsis from 2 independent cohorts[J]. *JAMA*, 2014, 312(1): 90-92.
- [3] SINGER M, DEUTSCHMAN C S, SEYMOUR C W, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3)[J]. *JAMA*, 2016, 315(8): 801-810.
- [4] JARCZAK D, KLUGE S, NIERHAUS A. Sepsis-pathophysiology and therapeutic concepts[J]. *Frontiers in Medicine*, 2021, 8(2): 302-324.
- [5] FENG M Y, WANG X S, ZHAO Z M, et al. Enhanced heart attack prediction using extreme gradient boosting[J]. *Journal of Theory and Practice of Engineering Science*, 2024, 4(4): 9-16.
- [6] KONG G L, LIN K, HU Y H. Using machine learning methods to predict in-hospital mortality of sepsis patients in the ICU[J]. *BMC Medical Informatics and Decision Making*, 2020, 20(3): 1-10.
- [7] LI Y, LIU Z Z. Deep learning-based prediction of in-hospital mortality for sepsis[J]. *Scientific Reports*, 2024, 14(1): 372-380.
- [8] DIAS A, GOMEZ V C, VIOLA L R, et al. Fever is associated with earlier antibiotic onset and reduced mortality in patients with sepsis admitted to the ICU[J]. *Scientific Reports*, 2021, 11(1): 49-54.
- [9] SHAH P K, GINESTRA J C, UNGAR L H, et al. A simulated prospective evaluation of a deep learning model for real-time prediction of clinical deterioration among ward patients[J]. *Critical Care Medicine*, 2021, 49(8): 1312-1321.
- [10] LAURITSEN S M, KALØR M E, KONGSGAARD E L, et al. Early detection of sepsis utilizing deep learning on electronic health record event sequences[J]. *Artificial Intelligence in Medicine*, 2020, 104(5): 820-831.
- [11] DE SILVA K, JÖNSSON D, DEMMER R T. A combined strategy of feature selection and machine learning to identify predictors of prediabetes[J]. *Journal of the American Medical Informatics Association*, 2020, 27(3): 396-406.
- [12] JOHNSON A, BULGARELLI L, POLLARD T, et al. MIMIC-IV clinical database demo (Version 2.2) [EB/OL]. (2023-01-31). <https://doi.org/10.13026/dp1f-ex47>.
- [13] 蔡德润, 李红燕. 基于度量学习的无监督域适应方法及其在死亡风险预测上的应用[J]. *计算机研究与发展*, 2022, 59(3): 674-682.
- [14] CAI Derun, LI Hongyan. A metric learning based unsupervised domain adaptation method with its application on mortality prediction[J]. *Journal of Computer Research and Development*, 2022, 59(3): 674-682.
- [15] KURSA M B, RUDNICKI W R. Feature selection with the Boruta package[J]. *Journal of Statistical Software*, 2010, 36(5): 1-13.
- [16] BOMMERT A, SUN X, BISCHL B, et al. Benchmark for filter methods for feature selection in high-dimensional classification data[J]. *Computational Statistics & Data Analysis*, 2020, 143(2): 19-38.
- [17] JOHANSSON P I, HENRIKSEN H H, KARVELSSON S T, et al. LASSO regression shows histidine and sphingosine 1 phosphate are linked to both sepsis mortality and endothelial damage[J]. *European Journal of Medical Research*, 2024, 29(1): 71-85.
- [18] LI Jianuo, BAO Wurigula, WANG Xuemei, et al. Estimating aboveground biomass of boreal forests in northern China using multiple data sets[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, 389(2): 189-210.
- [19] CHEN T Q, GUESTRIN C. XGBoost: A scalable tree boosting system[C]//*Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, California: ACM, 2016: 785-794.
- [20] BHATT D, PATEL C, TALSANIA H, et al. CNN variants for computer vision: History, architecture, application, challenges and future scope[J]. *Electronics*, 2021, 10(20): 2470-2498.
- [21] VAHDAT A, KAUTZ J. NVAE: A deep hierarchical variational autoencoder[J]. *Advances in Neural Information Processing Systems*, 2020, 33(1): 667-679.
- [22] YANG Jianbo, SHEN Kaiquan, ONG Chongjin, et al. Feature selection for MLP neural network: The use of random permutation of probabilistic outputs[J]. *IEEE Transactions on Neural Networks*, 2009, 20(12): 1911-1922.

- [22] 徐鹤, 郑群力, 谢作玲, 等. 基于知识表示向量的可解释深度学习模型及其疾病预测应用[J]. 数据采集与处理, 2023, 38(4): 777-792.
XU He, ZHENG Qunli, XIE Zuoling, et al. Interpretable deep learning model based on knowledge representation vectors and its application in disease prediction[J]. *Journal of Data Acquisition and Processing*, 2023, 38(4): 777-792.
- [23] 王明霞, 熊贇. 基于对比学习的疾病诊断预测算法[J]. 计算机科学, 2023, 50(7): 46-52.
WANG Mingxia, XIONG Yun. Disease diagnosis prediction algorithm based on contrastive learning[J]. *Computer Science*, 2023, 50(7): 46-52.
- [24] ROSENBLATT F. The perceptron: A probabilistic model for information storage and organization in the brain[J]. *Psychological Review*, 1958, 65(6): 386-408.
- [25] MCMAHAN B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data [C]//*Proceedings of 20th International Conference on Artificial Intelligence and Statistics*. Florida State: Machine Learning Research, 2017: 1273-1282.
- [26] 莫慧凌, 郑海峰, 高敏, 等. 基于联邦学习的多源异构数据融合算法[J]. 计算机研究与发展, 2022, 59(2): 478-487.
MO Huiling, ZHENG Haifeng, GAO Min, et al. Multi-source heterogeneous data fusion based on federated learning[J]. *Journal of Computer Research and Development*, 2022, 59(2): 478-487.
- [27] GOLDBERGER A L, AMARAL L A N, GLASS L, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals[J]. *Circulation*, 2000, 101(23): 215-220.
- [28] SAADATI Y, IMTEAJ A, AMINI M H. Vertical federated learning: Principles, applications, and future frontiers[C]//*Proceedings of Distributed Machine Learning and Computing*. Florida State: Springer International Publishing, 2024: 111-127.
- [29] CHEN P P, DONG W, WANG J L, et al. Interpretable clinical prediction via attention-based neural network[J]. *BMC Medical Informatics and Decision Making*, 2020, 20(1): 11-20.
- [30] RAJPURKAR P, YANG J, DASS N, et al. Evaluation of a machine learning model based on pretreatment symptoms and electroencephalographic features to predict outcomes of antidepressant treatment in adults with depression: A prespecified secondary analysis of a randomized clinical trial[J]. *JAMA Network Open*, 2020, 3(6): 653-660.
- [31] OTCHERE D A, GANAT T O A, GHOLAMI R, et al. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models[J]. *Journal of Petroleum Science and Engineering*, 2021, 200(2): 182-202.
- [32] SAYED M, RIAÑO D, VILLAR J. Novel criteria to classify ARDS severity using a machine learning approach[J]. *Critical Care*, 2021, 25(1): 1-9.
- [33] WANG Y B, LIU L Q, WANG C. Trends in using deep learning algorithms in biomedical prediction systems[J]. *Frontiers in Neuroscience*, 2023, 17(2): 351-383.
- [34] WALLACE M L, MENTCH L, WHEELER B J, et al. Use and misuse of random forest variable importance metrics in medicine: Demonstrations through incident stroke prediction[J]. *BMC Medical Research Methodology*, 2023, 23(1): 144-156.
- [35] CHUNDURU A, KISHORE A R, SASAPU B K, et al. Multi chronic disease prediction system using CNN and random forest[J]. *SN Computer Science*, 2024, 5(1): 157-170.
- [36] MAHOTO N A, SHAIKH A, SULAIMAN A, et al. A machine learning based data modeling for medical diagnosis[J]. *Biomedical Signal Processing and Control*, 2023, 81(5): 481-496.

作者简介:



文婷(1999-),女,硕士研究生,研究方向:临床预测, E-mail: s220603016@stu.cqupt.edu.cn。



余雷(1982-),通信作者,男,博士研究生,研究方向:急诊危重症患者的救治, E-mail: yulei@cqmu.edu.cn。



李腊全(1986-),通信作者,女,博士,副教授,研究方向:医学图像处理、医学数据分析和机器学习, E-mail: lilq@cqupt.edu.cn。

A Sepsis Mortality Risk Prediction Model Based on Multi-feature Federated Learning

WEN Ting¹, YU Lei^{2*}, LI Laquan^{1*}

(1. School of Mathematics and Statistics, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;

2. Emergency Department, The Second Affiliated Hospital of Chongqing Medical University, Chongqing 400010, China)

Abstract: Sepsis refers to a systemic inflammatory response resulting from infections, and it carries a high risk of mortality in intensive care settings. Existing predictive models often rely on extracting single feature subsets from a larger set, failing to fully utilize the complex interactions between feature subsets, known as structural mutual information. This limitation reduces prediction accuracy. Structural mutual information not only captures dependencies between features at the same level of granularity but also reveals complex relationships across different granularities, enabling more precise detection of subtle changes in a patient's condition. To address this limitation, this study presents a novel sepsis prognosis model that deeply explores the structural mutual information within electronic health records, significantly enhancing the accuracy of mortality risk predictions. Experimental results show that the proposed model achieves notable improvements in predictive accuracy, providing clinicians with more dependable mortality risk assessments and clearer decision-making support.

Highlights:

1. Propose SPPM, a sepsis mortality risk prediction model based on multi-feature joint learning, which exploits structural mutual information among clinical feature subsets to improve 30-day ICU mortality risk assessment.
2. Develop a multi-subset learning strategy that combines multiple feature selection algorithms with customized MLP predictors, enabling different clinical feature groups to capture complementary risk patterns from electronic medical records.
3. Integrate federated learning-based model aggregation to enhance prediction performance while reducing direct dependence on centralized raw patient data, providing a privacy-aware framework for clinical decision support.

Key words: deep learning; integrated learning; feature learning; sepsis prognosis; decision support

Foundation items: National Natural Science Foundation of China(Nos.61902046, 61901074, 62076044); China Postdoctoral Science Foundation(No.2021M693771); Natural Science Foundation of Chongqing (No.CSTB2022NSCQ-MSX0145).

Received: 2025-10-04; **Revised:** 2026-01-18

***Corresponding authors, E-mail:** yulei@cqmu.edu.cn; lilq@cqupt.edu.cn.